

Manuel Feißt

Dr. sc. hum.

Incorporation of Historical Two-Arm Data in Clinical Trials with Binary Outcome

Fach/Einrichtung: Medizinische Biometrie und Informatik

Doktorvater: Prof. Dr. Meinhard Kieser

Das Ziel dieser Arbeit war es, zu untersuchen, ob und wie Daten einer bereits durchgeführten (historischen) zweiarmligen klinischen Studie in eine neue klinische Studie eingebunden werden können. Es wurde überprüft, ob diese Einbindung mit einem Mehrwert im Sinne einer Erhöhung der Power beziehungsweise einer Reduzierung des erforderlichen Stichprobenumfangs einer neuen klinischen Studie im Vergleich zu einer konventionellen Studie ohne Einbindung historischer Daten einhergehen kann. Eine Reduzierung des erforderlichen Stichprobenumfangs reduziert in der Regel auch den zeitlichen Aufwand und die Kosten einer neuen klinischen Studie. Dies kann daher aus operativer Sicht als sehr wünschenswert angesehen werden. Darüber hinaus kann eine Reduzierung des Stichprobenumfangs und der Dauer einer klinischen Studie auch aus Sicht der Patienten als vorteilhaft angesehen werden, da wirksame Behandlungen schneller ihren Weg in die klinische Praxis finden können.

In einem regulatorischen Kontext ist eine notwendige Bedingung für die erfolgreiche Einbindung historischer Daten in eine neue Studie die Kontrolle der Wahrscheinlichkeit eines Fehlers 1. Art unterhalb eines vorgegebenen Signifikanzniveaus. Im Allgemeinen vergrößert sich jedoch die Wahrscheinlichkeit des Fehlers 1. Art mit steigendem Anteil an eingebundenen historischen Daten. Daher wurden in dieser Arbeit Ansätze entwickelt, die auf einer der sogenannten Power-Prior-Methode beruhen, welche es erlaubt, den Anteil der in die neue Studie einfließenden historischen Daten zu kontrollieren. Diese Bayes'sche Methode wurde in einen frequentistischen Rahmen überführt, da die statistischen Konzepte des Fehlers 1. Art und der Power ursprünglich innerhalb der Inferenztheorie eines frequentistischen Settings entwickelt wurden.

Im Rahmen dieser Arbeit wurde gezeigt, dass für ein zweiseitiges statistisches Testproblem mit steigendem Anteil an historischen Daten aus zwei Studienarmen die Wahrscheinlichkeit eines Fehlers 1. Art zunächst abnimmt, bevor er zunimmt. Dadurch war es möglich, bei gleichzeitiger Kontrolle der Wahrscheinlichkeit eines Fehlers 1. Art zum vorgegebenen Signifikanzniveau, einen entsprechenden Anteil an historischen Daten in eine neue Studie einzubinden. Es wurde gezeigt, dass das Ausmaß dieses Anteils von verschiedenen Parametern abhängt. Unter der Berücksichtigung dieser sogenannten Störparameter, wurden drei verschiedene Ansätze entwickelt um den Anteil der einzubeziehenden historischen Daten zu bestimmen. Im weiteren Verlauf dieser Arbeit wurden diese drei Ansätze insbesondere bezüglich der Möglichkeit Stichprobenumfang einzusparen untersucht und miteinander verglichen. Es konnte gezeigt werden, dass durch die Einbeziehung historischer Daten in vielen Szenarien die Power zur Aufdeckung des gleichen Effekts, wie er in den historischen Daten beobachtet wurde, erhöht werden kann. Folglich kann der erforderliche Stichprobenumfang für eine neue Studie für viele praktisch relevante Situationen reduziert werden. Es wurden jedoch auch einige Szenarien

identifiziert, in denen die Einbeziehung historischer Daten nicht mit einem Mehrwert verbunden ist.

Die in dieser Arbeit entwickelten Ansätze sind mit einem hohen Rechenaufwand verbunden. Es wurden daher praktische Empfehlungen gegeben, um diesen zu verringern. Darüber hinaus wurde ein Algorithmus für die Bestimmung des optimalen Stichprobenumfangs entwickelt, der den Rechenaufwand bei den entwickelten Verfahren deutlich reduziert.

Zusammenfassend wurde in dieser Arbeit gezeigt, dass die Einbeziehung historischer Daten aus zwei Studienarmen in eine neue Studie mit einem Mehrwert verbunden sein kann. Dieser Mehrwert spiegelt sich im Sinne eine Erhöhung der Power zugunsten des Effekts, wie er in den historischen Daten beobachtet wurde beziehungsweise in einer Reduzierung des erforderlichen Stichprobenumfangs wider. Gleichzeitig wird dabei die Wahrscheinlichkeit eines Fehlers 1. Art durch das vorgegebene Signifikanzniveau eingehalten. Die Existenz und das Ausmaß dieses Mehrwerts hängt jedoch maßgeblich von den zugrundeliegenden historischen Daten ab. Es wurden sowohl Szenarien identifiziert, die mit einem hohen als auch solche, die mit gar keinem Mehrwert einhergehen.